Please replace the paragraph beginning at page 2, line 11 with the following rewritten paragraph:

--Accordingly, in one aspect, the invention features a nucleic acid molecule that encodes a 14094 protein or polypeptide, e.g., a biologically active portion of the 14094 protein. In a preferred embodiment the isolated nucleic acid molecule encodes a polypeptide having the amino acid sequence of SEQ ID NO:2 or SEQ ID NO:12. In other embodiments, the invention provides isolated 14094 nucleic acid molecules having the nucleotide sequence shown in SEQ ID NO:1, SEQ ID NO:11, SEQ ID NO:3, or SEQ ID NO:13. In still other embodiments, the invention provides nucleic acid molecules that are substantially identical (e.g., naturally occurring allelic variants) to the nucleotide sequence shown in SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:11, or SEQ ID NO:13. In other embodiments, the invention provides a nucleic acid molecule which hybridizes under a stringency condition described herein to a nucleic acid molecule comprising the nucleotide sequence of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:11, or SEQ ID NO:13, wherein the nucleic acid encodes a full length 14094 protein or an active fragment thereof.--

B2

Please replace the paragraph beginning at page 3, line 11 with the following rewritten paragraph:

--In other embodiments, the invention provides 14094 polypeptides, e.g., a 14094 polypeptide having the amino acid sequence shown in SEQ ID NO:2, or SEQ ID NO:12; an amino acid sequence that is substantially identical to the amino acid sequence shown in SEQ ID NO:2, or SEQ ID NO:12; or an amino acid sequence encoded by a nucleic acid molecule having a nucleotide sequence which hybridizes under a stringency condition described herein to a nucleic acid molecule comprising the nucleotide sequence of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:11, or SEQ ID NO:13 or an active fragment thereof.--

B3

Please replace the paragraph beginning at page 8, line 18 with the following rewritten paragraph:

--*Figure 4* is a bar graph depicting the expression of 14094 RNA in a panel of normal and tumor human tissues, including breast, colon, liver, and lung, detected using

B4

B4/1 TAQMAN® analysis. 14094 RNA expression in normal (solid bars) and malignant ("diseased"; hatched bars) tissues from the breast, colon, liver and lung is shown. Elevated expression of 14094 RNA was detected in malignant tissues relative to normal tissues.--

Please replace the paragraph beginning at page 8, line 23 with the following rewritten paragraph:

B5 --*Figure 5* is a bar graph depicting the expression of 14094 RNA in a panel of normal and tumor human ovarian samples, detected using TAQMAN® analysis. Elevated expression of 14094 RNA was detected in malignant ovarian tissues relative to normal tissues.--

Please replace the paragraph beginning at page 8, line 27 with the following rewritten paragraph:

B6 --*Figure 6* is a bar graph depicting the expression of 14094 RNA in a panel of cell lines, detected using TAQMAN® analysis. Elevated expression of 14094 RNA was detected in DLD-1 and SW 620 cells lines. Both DLD-1 and SW620 are cell lines derived from colorectal carcinomas. SW620 is a lymph node metastasis of a colorectal carcinoma.--

Please replace the paragraph beginning at page 10, line 22 with the following rewritten paragraph:

B7 --For general information regarding PFAM identifiers, PS prefix and PF prefix domain identification numbers, refer to Sonnhammer et al. (1997) *Protein* 28:405-420.--

Please delete the paragraph beginning at page 10, line 25 and amend Table 1 as follows:

--

B8

Table 1: Summary of Sequence Information for 14094

Gene	cDNA	ORF	Polypeptide
14094	SEQ ID NO:1,	SEQ ID NO:3	SEQ ID NO:2
14094	SEQ ID NO:11	SEQ ID NO:13	SEQ ID NO: 12

--

Please replace the paragraph beginning at page 14, line 16 with the following rewritten paragraph:

--As used herein, the term "trypsin domain" (or a "trypsin-chymotrypsin" domain) refers to a protein domain having an amino acid sequence of from about 50 to about 350 amino acid residues and having a bit score for the alignment of the sequence to the trypsin domain (HMM) of at least 80. Preferably, a trypsin domain includes at least about 100 to about 300 amino acids, more preferably about 150 to about 250 amino acid residues, about 200 to about 230, or about 226 amino acids and has a bit score for the alignment of the sequence to the trypsin domain (HMM) of at least 100, preferably at least 200, more preferably at least 220, and most preferably 250 or greater. The trypsin domain (HMM) has been assigned the PFAM Accession (PF00089). An alignment of the trypsin domain (from about amino acids 217 to about 443 of SEQ ID NO:2) of human 14094 with a consensus amino acid sequence derived from a hidden Markov model (PFAM) is depicted in Fig. 3A. An alignment of the trypsin domain (from about amino acids 217 to about 443 of SEQ ID NO:2) of human 14094 with a consensus amino acid sequence derived from another hidden Markov model (SMART) is depicted in Fig. 3B.--

B9

Please replace the paragraph beginning at page 15, line 9 with the following rewritten paragraph:

--To identify the presence of a "trypsin" domain in a 14094 protein sequence, and make the determination that a polypeptide or protein of interest has a particular profile, the amino acid sequence of the protein can be searched against a database of HMMs (e.g., the Pfam database, release 2.1) using the default parameters. For example, the hmmsf program, which is available as part of the HMMER package of search programs, is a family specific default program for MILPAT0063 and a score of 15 is the default threshold score for determining a hit. Alternatively, the threshold score for determining a hit can be lowered (e.g., to 8 bits). A description of the Pfam database can be found in Sonhammer *et al.* (1997) *Proteins* 28(3):405-420 and a detailed description of HMMs can be found, for example, in Gribskov *et al.*(1990) *Meth. Enzymol.* 183:146-159; Gribskov *et al.*(1987) *Proc. Natl. Acad. Sci. USA* 84:4355-4358;

B10

B10

Krogh *et al.*(1994) *J. Mol. Biol.* 235:1501-1531; and Stultz *et al.*(1993) *Protein Sci.* 2:305-314, the contents of which are incorporated herein by reference. A search was performed against the PFAM HMM database resulting in the identification of a "trypsin domain" in the amino acid sequence of human 14094 at about residues 217 to about 443 of SEQ ID NO:2 with a bit score of 293 (see Figs. 1 and 3).--

Please replace the paragraph beginning at page 17, line 15 with the following rewritten paragraph:

B11

--As used herein, the term "scavenger receptor cysteine-rich domain" includes an amino acid sequence of about 80 to 120 amino acid residues in length and having a bit score for the alignment of the sequence to the scavenger receptor cysteine-rich domain (HMM) of at least 3. Preferably, a scavenger receptor cysteine-rich domain includes at least about 80 to 120 amino acids, more preferably about 87 to 110 amino acid residues, or about 90 to 100 amino acids and has a bit score for the alignment of the sequence to the scavenger receptor cysteine-rich domain (HMM) of at least 3 or greater. The scavenger receptor cysteine-rich domain (HMM) has been assigned the PFAM Accession Number PF00530. An alignment of the scavenger receptor cysteine-rich domain (amino acids 110 to 205 of SEQ ID NO:2) of human 14094 with a consensus amino acid sequence (SEQ ID NO:7) derived from a hidden Markov model is depicted in Figure 4B.--

Please replace the paragraph beginning at page 26, line 24 with the following rewritten paragraph:

B12

--The comparison of sequences and determination of percent identity between two sequences can be accomplished using a mathematical algorithm. In a preferred embodiment, the percent identity between two amino acid sequences is determined using the Needleman and Wunsch ((1970) *J. Mol. Biol.* 48:444-453) algorithm which has been incorporated into the GAP program in the GCG software package, using either a Blossum 62 matrix or a PAM250 matrix, and a gap weight of 16, 14, 12, 10, 8, 6, or 4 and a length weight of 1, 2, 3, 4, 5, or 6. In yet another preferred embodiment, the percent identity between two nucleotide sequences is determined using the GAP program in the GCG software package, using a NWSgapdna.CMP

matrix and a gap weight of 40, 50, 60, 70, or 80 and a length weight of 1, 2, 3, 4, 5, or 6. A particularly preferred set of parameters (and the one that should be used unless otherwise specified) are a Blossum 62 scoring matrix with a gap penalty of 12, a gap extend penalty of 4, and a frameshift gap penalty of 5.--

B12
1

Please replace the paragraph beginning at page 27, line 8 with the following rewritten paragraph:

-- The nucleic acid and protein sequences described herein can be used as a "query sequence" to perform a search against public databases to, for example, identify other family members or related sequences. Such searches can be performed using the NBLAST and XBLAST programs (version 2.0) of Altschul, *et al.* (1990) *J. Mol. Biol.* 215:403-10. BLAST nucleotide searches can be performed with the NBLAST program, score = 100, wordlength = 12 to obtain nucleotide sequences homologous to 14094 nucleic acid molecules of the invention. BLAST protein searches can be performed with the XBLAST program, score = 50, wordlength = 3 to obtain amino acid sequences homologous to 14094 protein molecules of the invention. To obtain gapped alignments for comparison purposes, Gapped BLAST can be utilized as described in Altschul *et al.*, (1997) *Nucleic Acids Res.* 25:3389-3402. When utilizing BLAST and Gapped BLAST programs, the default parameters of the respective programs (e.g., XBLAST and NBLAST) can be used. --

B13

Please replace the paragraph beginning at page 32, line 1 with the following rewritten paragraph:

B14

--In preferred embodiments, nucleic acids include a nucleotide sequence which is about 311, 400, 500, 600, 700, 800, 900, 1000, 1100, 1200, 1300, 1400, 1500, 1600, 1700, 1800, 1900, 2000, 2100, 2200, 2300, 2400, 2500, 2600, 2700, 2800 or 2900 nucleotides in length and hybridizes under stringent hybridization conditions to a nucleic acid molecule of SEQ ID NO:1, SEQ ID NO:11, SEQ ID NO:3, or SEQ ID NO:13.--

Please replace the paragraph beginning at page 33, line 32 with the following rewritten paragraph:

B15

--Moreover, nucleic acid molecules encoding other 14094 family members and, thus, which have a nucleotide sequence which differs from the 14094 sequences of SEQ ID NO:1, 11, 13, or 3, are intended to be within the scope of the invention.--

Please replace the paragraph beginning at page 115, line 17 with the following rewritten paragraph:

--<u>Example 2: Tissue Distribution of 14094 mRNA by TAQMAN® Analysis</u>

B16

Endogenous human 14094 gene expression was determined using the Perkin-Elmer/ABI 7700 Sequence Detection System which employs TAQMAN® technology. Briefly, TAQMAN® technology relies on standard RT-PCR with the addition of a third gene-specific oligonucleotide (referred to as a probe) which has a fluorescent dye coupled to its 5' end (typically 6-FAM) and a quenching dye at the 3' end (typically TAMRA). When the fluorescently tagged oligonucleotide is intact, the fluorescent signal from the 5' dye is quenched. As PCR proceeds, the 5' to 3' nucleolytic activity of Taq polymerase digests the labeled primer, producing a free nucleotide labeled with 6-FAM, which is now detected as a fluorescent signal. The PCR cycle where fluorescence is first released and detected is directly proportional to the starting amount of the gene of interest in the test sample, thus providing a quantitative measure of the initial template concentration. Samples can be internally controlled by the addition of a